



A Study on Relaxation Algorithms for Optimization Problems Arising in Data Mining

著者	伊豆永 洋一
発行年	2016
その他のタイトル	データマイニングで生じる最適化問題に対する緩和法の研究
学位授与大学	筑波大学 (University of Tsukuba)
学位授与年度	2015
報告番号	12102甲第7686号
URL	http://hdl.handle.net/2241/00143344

氏 名	伊豆永 洋一			
学 位 の 種 類	博 士 (工学)			
学 位 記 番 号	博 甲 第 7686 号			
学位授与年月日	平成 28 年 3 月 25 日			
学位授与の要件	学位規則第 4 条第 1 項該当			
審 査 研 究 科	システム情報工学研究科			
学位論文題目	A Study on Relaxation Algorithms for Optimization Problems			
	Arising in Data Mining (データマイニングで生じる最適化問題に対する緩和法の研究)			
主 査	筑波大学 教授	工学博士	吉瀬 章子	
副 査	筑波大学 教授	博士 (理学)	繁野 麻衣子	
副 査	筑波大学 准教授	博士 (学術)	八森 正泰	
副 査	筑波大学 准教授	博士 (情報理工学)	小林 佑輔	
副 査	山口大学 教授	博士 (経営学)	渋谷 綾子	
副 査	筑波大学 教授	工学博士	山本 芳嗣	

論 文 の 要 旨

情報化が進む現代社会において、日々刻々と輩出されるデータから意味を抽出する作業は、ますます重要な課題となっている。本論文は、このようなデータマイニングを行う上で発生する最適化問題のいくつかに着目し、その厳密解を求めることを前提として、それらに対する現実的な時間で求解可能な緩和問題を提案し、計算実験によってその有効性を検証している。

本論文で取り上げられている最適化問題は、モジュラリティ最大化問題、モジュラリティ密度最大化問題、ランキング問題の 3 つである。以下ではそれぞれの問題と、本論文でのアプローチならびに得られた成果を述べる。

あるグラフの所与の頂点分割に対して、モジュラリティは各分割内の頂点間の枝数と頂点の度数によって与えられる尺度関数であり、この値が高いほど、コミュニティをよく検出した分割であると言われる。モジュラリティ最大化問題は、この関数の最大値を与える頂点分割を求める問題であり、NP 困難であることが知られている。本論文ではまず、この問題を集合分割問題として 0-1 整数計画問題に帰着させている。この問題には集合分割の数に対応する指数本の制約が含まれているが、最適解において有効となる制約はごく少数である。本論文ではこの事実を利用して、線形緩和と切除平面法を組み合わせたアルゴリズム、ラグランジュ緩和と列生成法を組み合わせたアルゴリズムの、2 つの反復アルゴリズムを提案している。いずれのアルゴリズムでも、変数の釘づけルールなど、問題の構造を生かした工夫を加えることにより、1 反復で解く問題のサイズを抑え、計算時間を短縮させている。

コミュニティ検出の例題としてよく知られた6つのグラフを用いた計算機実験では、いずれのアルゴリズムでも、現実的な時間内で精度のよい解が得られている。

モジュラリティ尺度の欠点として、枝の数に影響を受けやすく、小さなコミュニティが検出されにくいことが指摘されている。これを克服する尺度として提案されているのが、各分割を構成する頂点の個数も評価に含めたモジュラリティ密度であり、モジュラリティ密度最大化問題はこの尺度を最大にする頂点分割を求める問題である。本論文では、この問題の最大値が既知であるならば、この問題を等価な0-1半正定値計画問題に帰着できることを示したのち、この問題を二重非負錐緩和し、この緩和問題の解の固有ベクトルから元の問題の解を抽出する動的計画アルゴリズムを提案している。7つの例題に対する計算機実験では、比較的小さな問題であれば、10秒以内に最適解が求められている。

ランキング問題は、データの多クラス分類問題の1つであり、評価指標が1つであるという特徴を持っている。本論文ではこの問題をサポートベクターマシンとして定式化し、この問題固有の最適解の性質を用いて双対問題を表現し、計算負荷を減らすための行・列生成アルゴリズムを提案している。予備的な計算機実験からは、提案アルゴリズムの有効性が確かめられた。

審 査 の 要 旨

【批評】

データマイニングの手法は多岐に渡っており、対象とするデータも年々大規模化している。本論文では3つの問題が取り上げられているが、これらを選択している理由は明確ではなく、また計算機実験で対象としている例題の大きさも実用の立場からは十分とはいえない。以上を考えると、本論文がデータマイニングの現場に与えるインパクトは限定的であるかも知れない。しかし、取り上げられた問題に対するモデル化、アルゴリズムの考え方は、最適化研究の立場からみて極めて正統的であり、特に「大規模だが規則性のある制約を持つ最適化問題に対して、緩和問題を提案し、制約と変数の両者の生成を繰り返して問題の解を求める」という一貫した思想に貫かれている。そのための独自性の高い工夫も随所に提案されており、この結果、本論文の成果は3編の査読付き学術論文として採択されている。以上の理論的な貢献により、本論文は、博士（工学）の学位に十分な水準に達していると判断する。

【最終試験の結果】

平成28年2月9日、システム情報工学研究科において、学位論文審査委員の全員出席のもと、著者に論文について説明を求め、関連事項につき質疑応答を行った。その結果、学位論文審査委員全員によって、合格と判定された。

【結論】

上記の学位論文審査ならびに最終試験の結果に基づき、著者は博士（工学）の学位を受けるに十分な資格を有するものと認める。